УДК 004

МОДЕЛЬ ПОРУШНИКА З НЕПОВНОЮ ІНФОРМАЦІЄЮ

О. В. Кіреєнко[1]

[1] *Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Фізико-технічний інститут*

Анотація

В цій роботі представлено основні підходи побудови моделі порушника в умовах невизначеності. Описано характеристики порушників, які можна визначити лише деяким діапазоном значень, випадковою величиною, параметричною залежністю від інших характеристик. Запропоновано сценарну модель порушника з неповною інформацією.

Ключові слова: невизначеність, модель порушника, сценарна модель, змістовна модель, математична модель впливу

Вступ

Розробка моделі порушника є одним з перших кроків при проектуванні захищених інформаційних систем. Незалежно від рівня деталізації моделі часто виникає проблема недостовірності або неповноти даних про порушників. Джерелами даної проблеми є неповнота статистичних даних та хибні припущення, що були зроблені на основі власного досвіду. Якщо організація зазнала збитків від деякої специфічної атаки, то в моделі порушників буде відведено більший пріоритет тим порушникам, які здатні подібну атаку повторити [1]. Для законного збору інформації про потенційних порушників підходять лише пасивні методи. Ми не можемо розраховувати на впровадження шпигунських програм в комп'ютери користувачів мережі. Найбільш ефективним пасивним методом є збір інформації, яку надає потенційний порушник. В межах організації – це аналіз трафіку, перевірка електронних листів, перегляд записів з камер спостереження, про які відомо потенційному порушнику. В глобальному плані – це дослідження пошукових запитів та активності в соціальних мережах. Недолік даного підходу очевидний – потенційний порушник знає, що за ним стежать. Це дозволяє потенційному порушнику спотворити зібрану інформацію, додавши до неї шум. Якщо профіль порушника формується на основі відсотку підозрілих дій, порушнику достатньо підвищити рівень дозволеної активності (або знизити рівень підозрілої), щоб профіль став неактуальним. Якщо замість відсотку враховуються абсолютні значення, порушник почне приховувати інформацію, використовуючи шифрування, проксі-сервери, а також залучати інших осіб, що знаходяться поза підозрою.

1. Типи невизначеності даних

Існує декілька типів невизначеності. При розробці моделі порушника необхідно передбачити коректну обробку даних, значення яких нам невідомі, відомі в деякому діапазоні, враховуються лише при виконанні деякої умови. Розглянемо характеристики порушників, які можуть бути відомі в деякому діапазоні. Мірою досвіду порушника є наступне співвідношення:

$$Exp = \frac{s_1 U_F + s_2 U_S}{d_1 R_F + d_2 R_S},$$

де U_F (Unique Failure) – кількість унікальних невдалих спроб взлому; U_S (Unique Success) – кількість унікальних успішних спроб взлому; R_F,R_S (Repeated Failure/Success) – повторювані невдалі та успішні взломи; $d_1{>}0$, $d_2{>}0$ (degradation) – коефіцієнти деградації досвіду; $s_1 \geq 1$, $s_2 \geq 1$ (sharing) – коефіцієнти ділення досвідом. Якщо зловмисник – script kiddy, його досвід буде нижче від деякого граничного значення X_0 (кількість унікальних невдач та унікальних успіхів обмежена можливостями шкідливого ПЗ, не унікальні успіхи та невдачі постійно накопичуються для будь-яких d_1,d_2), і оцінка рівня загрози від такого порушника зводиться до оцінки загрози шкідливого ПЗ, яким він користується. Для script kiddy модель порушника має бути максимально спрощеною, так як всі можливі сценарії впливу – стандартні. Для того, щоб оцінка досвіду порушника адекватно відображала його можливості, необхідно ввести такі обмеження:

1) порушник зберігає свої засоби атаки і вміння їх використовувати. Ця умова потрібна для оцінки всього арсеналу шкідливого ПЗ, зібраного порушником при взломі інших систем;
2) масову атаку рахуємо як одну. При цьому формула змінюється з урахуванням % успіху. Якщо при масовій атаці було вражено 40% цілей, то U_F, R_F збільшують не на одиницю, а на $0,6$, а U_S, R_S – на $0,4$ відповідно;
3) 123 тривалі або часто повторювані атаки (напр. перебір паролю) теж рахують як одну атаку. Кожна невдала спроба вгадати пароль не збільшує R_F. Якщо перебір паролю є єдиним способом атаки, ми можемо збільшувати R_F при $d_1 \to 0$

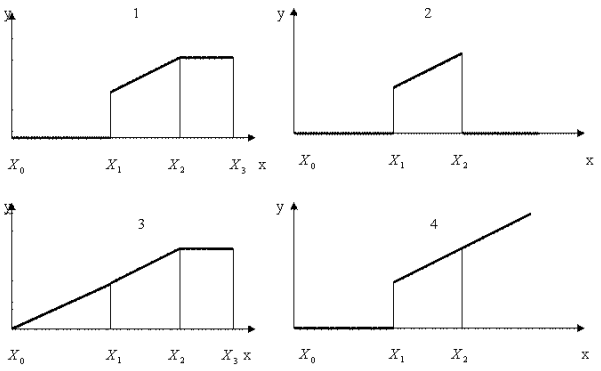

Рис. 1. Залежність рівня загрози зловмисника від характеристики в діапазоні $[X_0;X_3]$

4) якщо порушник – одинак, він навчається лише на власних невдачах та успіхах. $s_1 \approx 1$, $s_2 \approx 1$. Якщо порушник діє в команді $s_2 \approx 1$, $s_1{=}N$, де N – кількість членів команди.

При досягненні деякого критичного рівня досвіду X_2, порушник перестає отримувати досвід від атак нашої системи. Якщо на цьому етапі він не досяг успіху (кожна атака призводить до не унікальної невдачі), він буде змушений змінити ціль атаки.

При досягненні рівня досвіду X_1, вміння порушника переважують вплив везіння. Зниження досвіду відповідає деградації порушника. Знання порушника з часом втрачають актуальність. Якщо характеристика порушника є випадковою величиною, необхідно знайти її розподіл. Випадковою величиною може бути географічне розміщення порушника (або географічне розміщення вузла мережі, з якого здійснюють атаку). Якщо випадкова величина неперервна, її дискретизують. Інтервали дискретизації можуть відрізнятися. В межах інтервалу похідна функції не змінює знак. У випадку з географічним розміщенням порушника величина залишається незмінною в межах інтервалу. Якщо 50% порушників – інсайдери, 40% – порушники з нашої країни, а 10% – порушники із-за кордону, то співвідношення інтервалів буде 5:4:1, в межах кожного з цих інтервалів характеристика порушника не змінюється (похідна рівна нулю). Дані інтервали опрацьовуються аналогічно до попереднього випадку з $[X_1;X_2]$. Таким чином, невизначеність в межах діапазону значень є частковим випадком невизначеності випадкової величини, для якої на всіх інтервалах, крім одного, рівень загрози порушника або рівний нулю (наше припущення про порушника з досвідом нижче X_1), або не вищий від значення в точці X_2, для інтервалів справа від даного.

Рівень загрози порушника – це інтегральна оцінка його можливостей. До моделі порушника заносять лише ті характеристики, від яких залежить дана величина. Представлені на рис. 1 функції є функціями рівня загрози порушника (y) від його досвіду (x). Розподіл випадкової величини в межах діапазону може бути довільним.

На першому графіку представлено ситуацію, що відповідає загрозі державній таємниці. До точки X_1 рівень досвіду порушника недостатній для нанесення шкоди системі, а після точки X_2 рівень загрози не зростає. До точки X_1 функція тотожньо рівна нулю, так як на такому рівні везіння не впливає на результат. Успіх цілком залежить від перебору всіх можливих сценаріїв атаки із підвищенням їх складності. Так як до захисту державної таємниці висувають найжорсткіші вимоги, зростання досвіду після точки X_2 неможливе. На другому графіку показано аналогічну ситуацію, але тепер зловмисник з досвідом більшим ніж X_2 взагалі не зацікавлений атакувати нашу систему. Причиною відмови від атаки може бути її економічна недоцільність. Ще один можливий варіант – швидка деградація. Якщо зловмисник вибирає цілі із деякої множини, він має впорядкувати свої атаки так, щоб для кожної цілі його поточне значення досвіду X було не нижче, ніж значення X_1 відповідної цілі. Причиною швидкої деградації досвіду найчастіше стає випуск патчів та оновлень програмного забезпечення. В такому випадку порушник атакує системи відповідно до їх розкладу встановлення оновлень.

Хоча для всіх чотирьох графіків рівень досвіду лежить в діапазоні $[X_0;X_3]$, на другому графіку точку X_3 можна не позначати, так як нас цікавить лише інтервал $[X_1;X_2]$, в той час як на першому графіку нам бажано знати співвідношення інтервалів $[X_1;X_2]$ та $[X_2;X_3]$.

Графіки 3 та 4 – некоректні. Їх не можна використовувати для побудови моделей. На графіку 3 немає розриву в точці X_1. Це означає, що везіння залишається суттєвим і на ділянці $[X_1;X_2]$. В такому випадку ми не можемо виділити досвід порушника як окрему величину для моделі.

Графік 4 є некоректним, бо містить помилку екстраполяції. Загроза порушника не може безкінечно зростати. Кількість інтервалів завжди обмежена. Якщо рівень загрози продовжує зростати, треба обов'язково вказати точку X_3.

Останнім типом невизначеності, який розглянуто в цій статті, є параметрична залежність від інших характеристик. При такій невизначеності невідомим є не безпосереднє значення деякої характеристики зловмисника, а параметр, який використовується при її розрахунку. Прикладом подібної невизначеності є зацікавленість порушника-інсайдера, яка залежить від співвідношення заробітної плати в нашій організації та його гонораром від замовника [2],[3],[4]. Так як рівень оплати подібних зловмисних дій нам невідомий, ми повинні приблизно вказати його max та min значення. Значення невідомого нам параметра є випадковою величиною з даного діапазону. Таким чином, цей тип невизначеності є узагальненням попередніх. При такій невизначеності ми підставляємо в модель не саму випадкову величину, а функцію від неї. Попередні два випадки можна вважати частковими випадками даного, якщо використати ТРИВІАЛЬНУ функцію, тобто функцію, що повертає вхідні параметри.

2. Сценарна модель порушника

За рівнем деталізації моделі порушників поділяють на змістовні, сценарні та математичні [5]. Змістовні моделі є найменш деталізованими і використовуються для визначення мотивації порушників. Сценарні моделі дозволяють розглянути різні варіанти впливу кожного порушника на систему. Математичні моделі впливу використовують для кількісних оцінок збитків. Для побудови математичної моделі впливу необхідно мати чітке уявлення про порушників. Якщо інформація про порушників неповна та/або недостовірна, будувати математичну модель недоцільно, так як отримані кількісні оцінки будуть або суперечливими, або надто загальними. При побудові сценарної моделі, неповнота/недостовірність даних впливає лише на частину сценаріїв. Сценарна модель з неповною інформацією містить адекватні та необґрунтовані/превентивні сценарії впливу.

Адекватні сценарії впливу – це сценарії, що повністю обґрунтовані моделлю порушника. Адекватні сценарії впливу можна отримати в моделі з неповною інформацією в тому випадку, коли відсутня частина інформації не впливає на сам сценарій. Прикладом адекватного сценарію впливу є сценарій віддаленої атаки порушника на систему, при неповноті інформації про географічне розміщення порушника (в нашій країні чи із-за кордону). Для транснаціональних корпорацій географічне розміщення порушника не є суттєвим для даного сценарію, доки порушник знаходиться в «межах досяжності». Тобто, якщо в транснаціональної компанії є представництво в деякій країні, порушники з цієї країни не відрізняються в межах моделі від порушників із країни, в якій розміщено головний офіс корпорації.

Превентивні сценарії впливу – це сценарії, що були розглянуті без достатнього обґрунтування. Прикладом превентивного сценарію є зловмисні дії неблагонадійних за деякими критеріями (політичними/релігійними/расовими) працівників організації. Ядром всіх превентивних заходів є розглянуті організацією необгрунтовані сценарії.

До сценарної моделі не входять сценарії, які не є очевидними через неповноту/ недостовірність інформації (нам намагається нашкодити порушник, про існування якого нам невідомо), та сценарії, результат яких не можна передбачити.

Мінімізація помилок в сценарній моделі здійснюється за рахунок встановлення додаткових зв'язків між її елементами. Двофакторна сценарна модель порушника представляє повнозв'язний граф, вершинами якого є характеристики порушника. Ребра графу відповідають зв'язкам між характеристиками. В двофакторній моделі в явному вигляді вказано звязки (або відсутність звязків) для кожної пари характеристик. Складність і нелінійність зв'язків означає, що найкращим способом їх представлення є табличне представлення. Недоліком формульного представлення є виключення, які необхідно вказувати окремо. Так як діапазони значень деякої характеристики зловмисника, при яких зловмисник

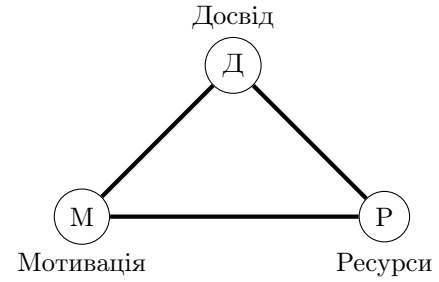

Рис. 2. Повнозв'язний граф для трьох характеристик

може нашкодити системі не співпадають в загальному випадку, їх можна скоротити. Замість початкових діапазонів використати їх перетин.

Розглянемо наступний приклад. На рис. 2 вказано зв'язки між трьома характеристиками. **Д**/**М**: З накопиченням досвіду мотивація порушника зростає (порушника стимулює безкарність у випадку невдачі, зростають шанси на успіх)[3]. Досвідчений порушник краще розуміє, як функціонує система, як її успішно атакувати, як ефективніше замітати сліди.

Д/**Р**: Досвідчений порушник діє на ринкову перспективу. Порушник, що використовує власні ресурси, зацікавлений не лише в успішній атаці, а й в реалізації добутого ресурсу. Витрати на проведення атаки можуть перевищувати потенційний прибуток, якщо порушник вмотивований не лише економічно. Витрати порушника ніколи не перевищують його бюджет. Порушник-виконавець може також використовувати ресурси, предоставлені замовником. Порушник-виконавець теж може проводити економічно недоцільні атаки, якщо вони узгоджені із замовником. Замовник знає кваліфікацію порушника і не виділяє кошти неперевіреному виконавцю. Порушник-виконавець використовує ресурси раціонально і не почне проведення атаки, якщо ресурсів недостатньо.

М/**Р**: Нехай A – мінімальна кількість ресурсів для проведення атаки. Порушник може провести атаку навіть якщо в нього ресурсів менше ніж A, але в такої атаки немає шансів на успіх. Величина B – максимальний можливий збиток системі від атаки. Навіть якщо систему продовжують атакувати, збитки не перевищать величину B. При збитках в розмірі B успішні атаки відтерміновують відновлення системи. При низьких значеннях мотивації і ресурсів порушник не стане об'єднуватися з іншими порушниками і просто відмовиться від атаки.

Зв'язок мотивації і ресурсів – ключовий в даній моделі. Саме цей зв'язок вказує на пріоритет порушника. У порушника може бути достатньо досвіду, ресурсів та вмотивованості для проведення атаки, але атака може не відбутися. Наприклад, у спецслужб є мотивація слідкувати за всім населенням, але слідкування за злочинцями, терористами, громадськими діячами має більший пріоритет, ніж слідкування за звичайними громадянами. Зв'язок **М**/**Р** вказує, чи вмотивований порушник витрачати ресурси саме на цю систему. Алгоритм застосування двофакторної сценарної моделі:

1) відкидаємо порушників, у яких недостатньо ресурсів для проведення атаки;
2) якщо в порушника не вистачає досвіду, він переходить в категорію замовника, а в модель записуємо порушника-виконавця;
3) відкидаємо порушників, що незацікавлені атакувати нашу систему (напр. партнери по бізнесу, спецслужби);
4) визначаємо рівень загрози порушників, у яких є більш пріоритетні цілі;
5) якщо система вже зазнала збитків в розмірі B, відкидаємо порушників, яким про це відомо;
6) яідкидаємо необґрунтовані атаки (чи є сенс знищувати носії/резервні копії інформації після її викрадення?);
7) всіх, кого відкинули на кроці 1) комбінуємо в команди так, щоб їх ресурси перевищували A;
8) для всіх команд з бюджетом більше A перевіряємо сумарний досвід. Якщо досвіду недостатньо – команда переходить в категорію замовника, а в модель записуємо порушника-виконавця;
9) Повторюємо кроки 3),4),5),6) для команд. В кроках 3) і 4) відкидаємо команди, в яких виникають протиріччя щодо доцільності атаки нашої системи.

Параметрична залежність від інших характеристик порушника залишається нерозв'язаною проблемою для даної моделі. Обійти дану проблему можна, якщо додати параметри до множини характеристик порушника, але тоді виникає дві проблеми: модель стає громіздкою (кількість зв'язків – це кількість ребер в повнозв'язному графі, до якого ми додали вершин); модель перестає бути виключно моделлю порушника (сюди інтегруються дані, що не стосуються порушників) [6].

Висновки

Ключовими проблемами при розробці моделі порушника є збір даних, їх обробка та перевірка. Основною функцією моделі порушника з неповною інформацією є обробка даних. Для адекватних сценаріїв впливу ми можемо відновити другорядну інформацію. Дослідження превентивних сценаріїв впливу дозволяє відновити важливі для моделі дані з імовірністю менше 100%. Якщо при обробці даних в моделі з'являються суперечності, це свідчить про некоректність даних. Обробка даних в адекватних сценаріях впливу дозволяє виявити причину суперечності. В превентивних сценаріях впливу суперечність можна виявити в тому випадку, коли не існує комбінації невідомих для сценарію параметрів, при якій суперечність відсутня. При неповній інформації розробка математичної моделі впливу може бути недоцільною. Сценарна модель впливу може предоставляти адекватні сценарії впливу навіть в умовах невизначеності. Необгрунтовані сценарії впливу є основою для розробки превентивних засобів захисту. Превентивний сценарій впливу може стати адекватним при уточненні даних.

На даний момент спостерігається тенденція до інтеграції моделі порушника в більш комплексні моделі [7], [8]. Це проявляється в поширенні моделі на всі типи порушників, а не тільки на порушників інформаційної безпеки; інтегруванні моделей порушника в системи виявлення вторгнень [9] та в інформаційні системи в цілому (модель модифікується у відповідь на будь-які дії користувачів системи).

Перелік використаних джерел

1. Human Factors and Information Security: Individual, Culture and Security Environment : Technical Report / Command, Control, Communications and Intelligence Division ; Executor: Kathryn Parsons, Agata McCormac, Marcus Butavicius, Lael Ferguson. — Edinburgh South Australia 5111 Australia : Defence Science and Technology Organisation, 2010. — 45 p.
2. Kim Sang Hoon, Yang Kyung Hoon, Park Sunyoung. An Integrative Behavioral Model of Information Security Policy Compliance // The Scientific World Journal. — 2014. — 12 p.
3. Bulgurcu Burcu, Cavusoglu Hasan, Benbasat Izak. Information security policy compliance: an empirical study of rationality-based beliefs and information security awareness // MIS Quarterly. — 2010. — September. — Vol. 34, no. 3. — P. 523–548.
4. Vance Anthony, Siponen Mikko. IS Security Policy Violations:A Rational Choice Perspective // Journal of Organizational and End User Computing. — 2012. — January-March. — Vol. 1, no. 24. — P. 21–41.
5. Полехіна Ю.М., Тимофєєв Д.С. Модель порушника. Мета та принципи розробки // Эффективные инструменты современных наук - 2010. — 2010. — Т. 1, № 4.
6. НД ТЗИ 3.7-002-99. — 1999. — С. 31.
7. Головань С.М. Вимоги до побудови моделі загроз інформаційних систем // Інформаційна безпека. — 2009. — Т. 2, № 2. — С. 77–84.
8. Maintenance of information security of the russian banking system organisations [Electronic resource]. — 2014. — June. — Access mode: `https://www.cbr.ru/Eng/analytics/Gubzi_docs_en/st-10-14_en.pdf`.
9. Detecting Threatening Behavior Using Bayesian Networks / Kathryn Laskey, Ghazi Alghamdi, Xun Wang et al. // Conference on Bihevioral Representation in Modeling and Simulation. — 2004. — P. 10.